

Лабораторный практикум " Установка и настройка кластера "
(на примере учебно-практических задач)

Содержание

Общее описание лабораторного практикума 3
Лабораторная работа 1. Установка и настройка операционной системы. 7
Лабораторная работа 2. Объединение узлов в кластер 10
Лабораторная работа 3. Создание высокопроизводительной системы 12
Лабораторная работа 4. Система пакетной обработки 15
Лабораторная работа 5. Исследование производительности 17
Литература 19

Общее описание лабораторного практикума

Аннотация:

Основная задача практикума - познакомить слушателей с современными технологиями, популярными подходами, а также некоторыми известными алгоритмами на основе которых создаются кластерные вычислительные системы. Предполагается, что практикум будет интересен как пользователям кластерных вычислительных систем, так и администраторам вычислительных кластеров построенных на базе UNIX серверов. Практикум может быть полезен в том числе, для лиц принимающих решения об аппаратной и программной составляющей создаваемого кластера. Изложенные в нем подходы могут быть использованы как отправная точка в научных исследованиях студентов в области сетевых технологий, операционных систем, систем поддержки параллельных и распределённых вычислений, распределённых файловых систем.

Цели и задачи практикума:

- Рассмотреть основные элементы вычислительной системы с распределенной памятью, используя как учебный пример задачу о развертывании вычислительного кластера;
- Выполнить ряд экспериментов по анализу производительности многопроцессорной вычислительной системы (кластера), показать узкие места в производительности и сравнить по результатам экспериментов эффективность различных коммуникационных библиотек;
- Подготовить по результатам лабораторного практикума сравнительную характеристику различных версий и платформ MPI;
- Получить представление об уровне предварительной подготовки обучаемых в области системного администрирования для уточнения плана дальнейшего обучения.

План выполнения лабораторного практикума:

Практикум включает:

1. Проведение практического занятия (2 часа) по установке и настройке операционной системы узлов кластера, настройке сетевых сервисов и сетевой файловой системы.
2. Проведение практического занятия (2 часа) по объединению узлов в кластер с помощью установки коммуникационной библиотеки MPI, библиотек линейной алгебры BLAS и LAPACK.
3. Проведение практического занятия (2 часа) по установке и настройке коммуникационной библиотеки BLACS и библиотеки линейной алгебры для распределенных систем ScaLAPACK.
4. Проведение практического занятия (2 часа) по установке системы пакетной обработки и мониторинга.
5. Проведение практического занятия (2 часа) по измерению производительности кластера на тесте Linpack, и производительности локальной сети.

В качестве задания повышенной сложности могут быть поставлены задачи по подбору параметров ядра, для оптимизации производительности; сетевого развертывания кластера; разделение сети на две части: сеть управления и сеть коммуникации; проведение сравнительного анализа сетевых файловых систем.

Слушатели выполняют настройку своего узла кластера, при этом, им необходимо учитывать то, что их узел будет работать в составе общего кластера, таким образом, их действия должны быть согласованы друг с другом. По завершении создания кластера слушатели выполняют несколько вычислительных экспериментов по оценке производительности полученной системы. По результатам работ проводится общая дискуссия.

Необходимое оборудование и ПО:

Вычислительный кластер лаборатории параллельных систем: четыре узла с 4-х ядерными процессорами с Gigabit сетевым интерфейсом объединенные в локальную сеть посредством неблокирующего коммутатора. Желательно наличие вторых сетевых интерфейсов у каждого узла и и второй коммутатор.

Выделенный NFS сервер.

Предполагается неограниченный доступ в Интернет.

В качестве несущей операционной системы предлагается Linux Ubuntu Server Edition.

Необходимый уровень подготовки:

Для успешного выполнения практикума предполагается наличие следующих знаний и практических навыков:

- Умение работать в командной оболочке bash.
- Базовая практика по разработке приложений на языке C в Unix.
- Начальные сведения по установке программ в Unix.
- Способность читать по английски.
- Владение поиском в Интернет

Методические рекомендации по выполнению:

Весь необходимый при выполнении практикума материал содержится в Интернет, для получения необходимой информации, достаточно корректно поставить поисковый запрос в большинстве из информационно-поисковых систем, например Google. Необходимый материал по установке Ubuntu, SSH, NFS, MPI, библиотекам BLAS, LAPACK, BLACS, ScaLAPACK может быть получен в Интернет (в частности, [1-3]). Полезная дополнительная информация может быть найдена в пособии [4].

Выполнение заданий практикума рекомендуется проводить группами из 2 человек. Первоначально все обучаемые, выполняют разработку базового

варианта работы. Далее после обсуждения результатов разработки каждая группа выбирает индивидуальный вариант постановки задания для продолжения работ.

Планируемый результат:

- Действующий кластер,
- Результаты выполненных вычислительных экспериментов,
- Таблицы и графики вычислительной производительности системы, латентности и пропускной способности сети полученные в результате выполненных экспериментов.

Форма отчета:

По одному из выполненных заданий группа готовит отчет, включающий:

- Описание поставленной задачи,
- Описание метода решения задачи,
- Результаты выполненных экспериментов (в форме таблиц и графиков),
- Выводы по полученным результатам.

Лабораторная работа 1. Установка и настройка операционной системы.

Цель:

- Получить практические навыки установки и настройки операционной системы Linux,
- Выполнить вычислительные эксперименты на кластере и оценить эффективность выполняемых параллельных вычислений.

Необходимый теоретический материал

- Общая структура вычислительного комплекса. Архитектура кластера типа Beowulf. Аппаратные технологии и устройства для организации коммуникаций в кластерах
- Инструментарий UNIX для администратора и продвинутого пользователя
- Файловые сервера в вычислительном центре, аппаратура используемая для создания файловых серверов.
- Протокол SSH, как протокол организации удалённого доступа, доступ по открытым ключам. SSH клиенты в различных операционных системах.
- Программное обеспечение высокопроизводительных систем.

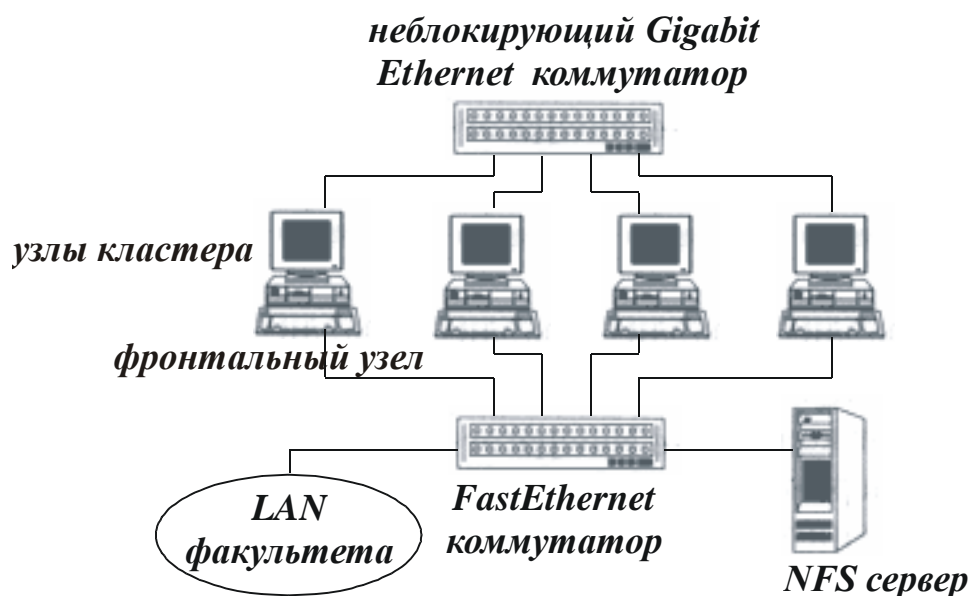


Рис.1 Архитектура Beowulf кластера

Задание 1. Установка Ubuntu Linux

Создать загрузочный диск для сетевой установки Ubuntu Server Edition.

Установить операционную систему в минимальной комплектации с одного из FTP зеркал.

С помощью менеджера пакетов установить SSH, если он не был установлен во время установки операционной системы.

Настроить сетевые интерфейсы узлов для работы в общей локальной сети

Настроить SSH для аутентификации без пароля с использованием RSA-ключей, на всех узлах кластера.

Настроить подключение к выделенному NFS серверу.

Задание 2. Установка полезных утилит

Установить текстовый редактор (nano, Emacs).

Установить пакет утилит поиска файлов findutils.

Установить Python, NumPy, SciPy

Дополнительное задание. Организация второй вспомогательной подсети

Для выполнения задания необходимо наличие двух сетевых интерфейсов на каждом узле, один из них может быть FastEthernet и второй коммутатор.

Настроить сетевые интерфейсы узлов и маршрутизацию таким образом, чтобы узлы будущего кластера взаимодействовали друг с другом по высокоскоростному каналу в одной подсети, а их взаимодействие с NFS сервером и внешним миром осуществлялось в другой.

Планируемый результат:

1. Подготовленная для организации кластера вычислительная инфраструктура состоящая из N узлов, с установленной на каждом узле ОС семейства Linux, объединенных в локальную сеть посредством неблокирующего коммутатора. Узлы подключены к выделенному NFS серверу. Узлы имеют возможность удаленного доступа друг к другу посредством SSH.

Лабораторная работа 2. Объединение узлов в кластер

Цель:

- Получить навыки установки и использования MPI,
- Выполнить вычислительные эксперименты на кластере и оценить эффективность выполняемых параллельных вычислений.

Необходимый теоретический материал

- Коммуникационные библиотеки для систем с распределенной памятью используемые в составе кластера. Программные средства организации взаимодействия процессов разнесённых на разные процессоры: MPI
- Организация параллельных вычислений для систем с распределенной памятью.
- Компиляторы используемые на многопроцессорных вычислительных системах.
- Сборка программных пакетов

Задание 1: Установка компиляторов и математических библиотек

Установить компиляторы языков C и FORTRAN. В качестве устанавливаемых компиляторов предлагается использование компиляторов Intel свободно распространяемых для Linux систем. Установить пакеты BLAS и LAPACK. Предлагается также использовать инструмент Intel, такой как MKL.

Установить NumPy, SciPy

С помощью менеджера пакетов установить интерпретатор скриптового языка Tcl.

Скачать и установить пакет динамической настройки окружения modules.

Задание 2: Установка MPI

С одного из зеркал скачать пакеты MPI-1 и MPI-2. Здесь предлагается использовать пакет mpich. Собрать mpich и mpich2 с разными устройствами (device). Установить и проверить установку с помощью примеров входящих в состав источника кода MPI. Сконфигурировать modules для поддержки возможности использования различных, проявившихся в системе, версий MPI. Настроить конфигурацию MPI так, чтобы все узлы строящегося кластера вошли в его состав.

Дополнительное задание. Сборка и запуск тестовых MPI приложений

Собрать и повести пробные запуски примеров MPI программ.

Провести тест Whetstone.

Планируемый результат:

1. Среда разработки.
2. Кластер MPI.

Лабораторная работа 3. Создание высокопроизводительной системы

Цель:

- Получить навыки установки и использования ScaLAPACK.
- Выполнить вычислительные эксперименты на кластере используя ScaLAPACK.

Необходимый теоретический материал

- Библиотеки линейной алгебры для высокопроизводительных систем.
- Организация параллельных вычислений для систем с распределенной памятью.
- Компиляторы используемые на многопроцессорных вычислительных системах.
- Сборка программных пакетов

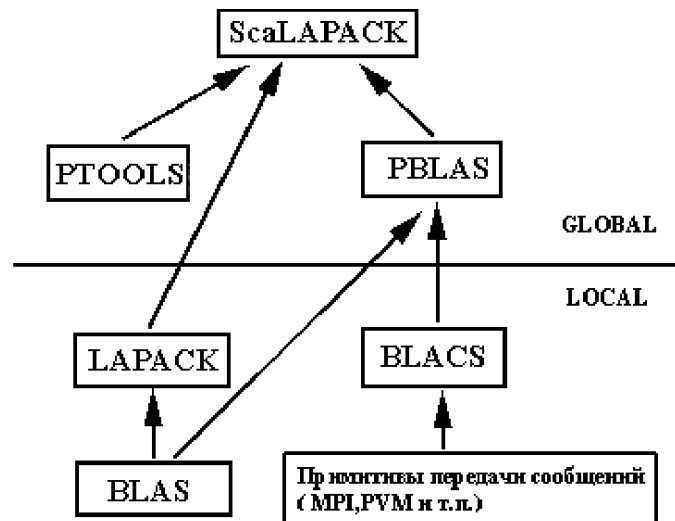


Рис.2 Структура ScaLAPACK

Задание 1: Сборка и настройка коммуникационной библиотеки BLACS

Скачать и собрать BLACS с использованием различных версий MPI существующих в системе. Сконфигурировать modules для поддержки возможности использования различных, проявившихся в системе, версий BLACS.

Задание 2: Сборка библиотеки линейной алгебры для распределенных систем ScaLAPACK.

Скачать и собрать ScaLAPACK.

Проверить работоспособность запуском примеров.

Дополнительное задание. Сборка и запуск тестовых задач ScaLAPACK

Собрать тесты ScaLAPACK.

Провести тест решения систем линейных уравнений на кластере.

Провести тест решения задачи нахождения собственных значений на кластере.

Планируемый результат:

1. Система для высокопроизводительных вычислений..
2. Таблицы и графики результатов выполненных вычислительных экспериментов.

Лабораторная работа 4. Система пакетной обработки

Цель:

- Получить навыки установки и использования системы пакетной обработки,
- Выполнить вычислительные эксперименты на кластере путем постановки задач в очередь.

Необходимый теоретический материал

- Организация параллельных вычислений для систем с распределенной памятью.
- Компиляторы используемые на многопроцессорных вычислительных системах.
- Сборка программных пакетов
- Организация системы очередей задач пользователей. Наиболее известные системы ведения очередей задач пользователей.

Задание 1: Установка системы пакетной обработки.

Собрать и настроить Torque.

Сделать пробные запуски задач.

.

Задание 2: Установка системы мониторинга кластера.

Собрать и настроить Ganglia.

Отследить работу запущенных задач и работу системы в целом.

Дополнительное задание. Запуск заданий

Запустить продолжительные (>2 часов) задачи.

Оценить работу системы очередей.

Планируемый результат:

1. Вычислительная система.
2. Таблицы и графики результатов выполненных вычислительных экспериментов.

Лабораторная работа 5. Исследование производительности

Цель:

- Получить навыки проведения теста Linpack,
- Провести вычислительные эксперименты на кластере для оценки факторов влияющих на производительность.

Необходимый теоретический материал

- Linpack benchmark
- Организация параллельных вычислений для систем с распределенной памятью.
- Сборка программных пакетов

Задание 1: Тест Linpack

Собрать Linpack HPC benchmark.

Провести тестирование производительности собранного кластера

Найти оптимальные параметры запуска.

.

Задание 2: Тест производительности сети

Собрать тест производительности сети.

Провести тестирование производительности сети.

Сделать выводы.

Дополнительное задание. Сравнение производительности

Собрать тесты Linpack с использованием различных версий MPI и BLAS.

Сделать анализ производительности.

Планируемый результат:

1. Результаты тестов
2. Оптимизированный кластер..
3. Таблицы и графики результатов выполненных вычислительных экспериментов.

Литература

1. Ubuntu - <http://www.ubuntu.com/>
2. Netlib Repository - <http://www.netlib.org/>
3. Информационные материалы по MPI -
<http://www.mpi-forum.org>
4. Лацис А.О. Как построить и использовать суперкомпьютер. -М.: Бестселлер, 2003.
5. Таненбаум Э., М. Стеен. Распределенные системы: принципы и парадигмы. СПб.:Питер, 2005, 878 с.
6. Корнеев В.В. Параллельные вычислительные системы. - М. «Нолидж», 1999.
7. Воеводин В.В., Воеводин Вл.В. Параллельные вычисления. -СПб: BHV, 2002.
8. Jack J. Dongarra, Performance of Various Computers Using Standard Linear Equations Software -
<http://www.netlib.org/benchmark/performance.ps>
9. The ScaLAPACK Project. – <http://www.netlib.org/scalapack>
- 10.ScaLAPACK Users' Guide. 1997. –
http://www.netlib.org/scalapack/scalapack_home.html
http://rsusu1.rnd.runnet.ru/ncube/scalapack/scalapack_home.html
- 11.Кластер (практическое руководство). - <http://cluster.linux-ekb.info/>
- 12.MPI: A Message-Passing Interface Standard. Message Passing Interface Forum. – Version 1.1. 1995. – <http://www-unix.mcs.anl.gov/mpi>
- 13.Portable Batch System. – <http://www.openpbs.org>
- 14.MPI: The Message Passing Interface. -
http://parallel.ru/tech/tech_dev/mpi.html
- 15.Официальный сайт TORQUE — <http://clusterresources.com>

16. Ganglia Monitoring System - <http://ganglia.info/>
17. Информационно-аналитический центр по параллельным вычислениям - <http://parallel.ru>
18. А.А. Букатов, В.Н. Дацюк, А.И. Жегуло Программирование многопроцессорных вычислительных систем. - Ротов-на-Дону: Издательство ООО «ЦВВР», 2002.